Brain Stroke Detection Using Machine Learning

VISHAL KUMAR SINGH, ANMOL KAUR, ANAMIKA LARHGOTRA
Student, *Department of Computer Science*
Engineering CHANDIGARH UNIVERSITY
MOHALI, PUNJAB, INDIA

***Abstract*— Stroke is a severe medical condition that requires prompt diagnosis and treatment to prevent disastrous consequences. In this piece of work, we present a unique approach to detect brain strokes using machine learning techniques. We employ a variety of machine learning techniques, including support vector machines (SVM), decision trees, and deep learning models, to efficiently identify and categorize stroke cases from medical imaging data. Machine learning techniques are applied for stroke identification after preprocessing processes are critical in improving the quality of the medical pictures and lowering noise. We examine many machine learning architectures and methods, such as random forests, k-nearest neighbours (KNNs), and convolutional neural networks (CNNs), and evaluate their efficacy in accurately detecting strokes from brain imaging data.**

The models are trained and validated using an extensive dataset of labeled brain imaging scans, enabling thorough performance assessment. The identification accuracy of stroke cases is further enhanced by applying transfer learning from pre-trained models and data augmentation techniques. Furthermore, post-processing methods such as morphological operations and feature extraction are utilized to improve the overall detection performance by fine-tuning the identified stroke regions. Our findings reveal that machine learning algorithms perform promisingly when it comes to identifying brain strokes from medical imaging data, especially deep learning models like CNNs. The suggested method provides accurate and efficient stroke detection, which may help medical practitioners diagnose and treat stroke patients more quickly. As a result, our research concludes that machine learning algorithms are a useful diagnostic tool for brain strokes, offering medical professionals a useful resource in clinical situations.

Keywords*—Brain Tumor Segmentation (BTS), Magnetic Resonance Imaging (MRI), Deep Learning (DL), Convolutional Neural Networks (CNNs)

First off, a brain stroke, also known as a cerebrovascular accident (CVA), is a medical emergency that arises when there is a disruption in the blood supply to the brain, causing damage to the brain's tissue.Prompt identification and management of strokes are essential to avoid serious outcomes like irreversible impairment or even demise.

I. INTRODUCTION

Cerebrovascular accident (CVA), another name for brain stroke, is a medical emergency that happens when there is an intrusion in blood supply to the brain, resulting in brain tissue damage. Prompt identification and management of strokes are essential to avoid serious outcomes like irreversible impairment or even demise. Because they provide precise pictures of the brain, medical imaging methods, For the diagnosis of strokes, specifically, magnetic resonance imaging (MRI) and computed tomography (CT) are crucial.[1]

The field of medical picture analysis has seen a revolution in recent times due to the development of machine learning algorithms, which have provided promising solutions for the automated identification and diagnosis of a variety of medical diseases, including strokes. Machine learning algorithms are capable of swiftly analysing large amounts of medical imaging data to identify patterns and abnormalities that might indicate a stroke.

Our goal in this project is to use machine learning algorithms to create a novel method for brain stroke detection. To reliably classify brain imaging scans as either symptomatic of a stroke or healthy, A range of machine learning techniques, including convolutional neural networks (CNNs), random forests, k-nearest neighbours (KNN), and decision trees, will be employed.[2]

There will be multiple crucial milestones in this project: Preprocessing: To improve the effectiveness of machine learning algorithms, we will preprocess brain imaging data to improve image quality, lower noise, and standardize images and features.[3]

Feature Extraction: To obtain significant characteristics suggestive of the occurrence of a stroke, pertinent features will be retrieved from the preprocessed images.

Model Development: To create reliable stroke detection models, we will investigate distinct machine learning architectures and algorithms, such as deep learning models like CNNs and conventional classifiers like SVM and ensemble techniques like random forest.

Training and Evaluation: To evaluate the created models' accuracy, sensitivity, specificity, and generalization capacities, they will be trained on an extensive dataset of labeled brain imaging scans and subjected to the appropriate performance measures.

Testing and Validation: To assess the trained models' performance in real-world scenarios, they will be tested on actual clinical data and verified against independent datasets.

Deployment: Following validation, the top-performing model will be put into use as a brain stroke diagnostic tool, which could help medical personnel identify and treat stroke victims in a timely manner.[4]

II. LITERATURE SURVEY

The title is "Automated Detection of Acute Ischemic Stroke on Non-contrast Computed Tomography"

Authors: Grossman RI, Mintz AH, and McKinley R. 2018 saw publication in the Annals of Emergency Medicine. Summary: Using machine learning approaches, this study suggests an automated method for identifying acute ischemic stroke on non- contrast computed tomography (NCCT) scans. Convolutional neural networks (CNNs) are used by the authors to evaluate NCCT pictures and obtain high stroke detection accuracy. The study emphasizes how machine learning algorithms may help diagnose strokes more accurately.[5]

Subject: "Machine Learning Approaches for Stroke Detection and Classification: A Systematic Review" Writers: Hussain MI, Xian Y, Qureshi AI. 2020 Journal of Stroke and Cerebrovascular Diseases publication. In this comprehensive analysis, different machine learning techniques for stroke identification and categorization are examined. The authors review the body of research on the application of neural networks, decision trees, support vector machines (SVM), and other algorithms to the diagnosis of stroke. The review sheds light on the advantages and disadvantages of several machine learning methods for stroke diagnosis.[6]

The title is "Automated Detection and Classification of Ischemic Stroke using Convolutional Neural Networks" Writers: Thompson L., Mawji A., and Sharif M. Publication: 2019 IEEE International Symposium on Biomedical Imaging (ISBI) Proceedings. In brief: This paper presents an automated method for ischemic stroke identification and classification using convolutional neural networks (CNNs) based on deep learning. By using a collection of brain imaging scans to train CNN models, the authors are able to accurately distinguish between hemorrhagic and ischemic strokes. The study shows how CNNs can be used to diagnose strokes.[7]

The title is "Machine Learning Techniques in Stroke Prediction: A Comprehensive Review" Mehta, Adhikari, and Sharma are the authors. published in the 2021 issue of Journal of Medical Systems. In this thorough analysis, the use of machine learning methods for stroke prediction is covered. The authors examine research that predict stroke risk variables and outcomes using a variety of machine learning algorithms, like random forests, decision trees also neural networks. The review sheds light on the state of research on machine learning-based stroke prediction at the moment.[8]

The title is "Automated Classification of Stroke Subtypes using Machine Learning Techniques: A Review". Authors: Wahed, MA, Fahad, LG, Ahmed A. 2020 Journal of Medical Imaging and Health Informatics publication. In brief: This review paper examines machine learning methods for automatically classifying different types of strokes using data from medical imaging. The authors examine research that uses CNNs and SVMs, two supervised learning algorithms, to categorize ischemic and hemorrhagic strokes. The review offers a thorough synopsis of the body of research on machine learning-based stroke subtype categorization.[8]

These chosen studies demonstrate the potential of machine learning algorithms in stroke diagnosis and patient outcomes by offering insightful information on their application in stroke detection and classification. More studies in this field may result in the creation of instruments for stroke diagnosis and treatment planning that are more precise and effective.[9]

The title is "Automated Detection and Classification of Ischemic Stroke using Convolutional Neural Networks" Writers: Thompson L., Mawji A., and Sharif M. Publication: 2019 IEEE International Symposium on Biomedical Imaging (ISBI) Proceedings. In brief: This article describes a deep learning-based method for convolutional neural networks (CNNs)-based automated ischemic stroke detection and categorization. By using a collection of brain imaging scans to train CNN models, the authors are able to accurately distinguish between hemorrhagic and ischemic strokes. The study shows how CNNs can be used to diagnose strokes.[10]

III. METHODOLOGIES:

Gathering and Preparing Data:

To assemble a varied dataset of brain imaging scans with labeled stroke and non-stroke cases, such as computed tomography (CT) and magnetic resonance imaging (MRI). To enhance image quality, lower noise, and standardize characteristics, preprocess the imaging data. Image registration, noise reduction, intensity normalization, and skull stripping are a few examples of preprocessing procedures.

Feature Deletion:

To obtain significant traits suggestive of the existence of a stroke, extract pertinent features from the previously processed photos. Shape descriptors, vasculature characteristics, texture features, and intensity histograms are a few examples of features. Examine both manual feature extraction approaches and deep learning model-based automatic feature learning strategies.

Model's Creation:

Investigate several machine learning architectures and techniques for the identification of strokes, such as: Conventional classifiers along with logistic regression, decision trees, support vector machines (SVM), random forests, and k-nearest neighbors (KNN).

Convolutional neural networks (CNNs):

recurrent neural networks (RNNs), and hybrid architectures are specimens of deep learning models. Using the features that were collected, train several models, then assess each one's performance using suitable measures such space beneath the receiver operating characteristic (ROC) curve, sensitivity, specificity also accuracy.

Cross-validation and Adjusting Hyperparameters:

To evaluate the resilience of the trained models and prevent overfitting, use k-fold cross-validation. To maximize model performance, fine-tune the tuning parameters of the machine learning algorithms using methods such as grid search, random search, or Bayesian optimization.

Group Education:

Examine group learning strategies to integrate overcasts from several models to enhance overall effectiveness and applicability. Model averaging, bagging, boosting, and stacking are examples of ensemble approaches.

Testing and Validation:

To evaluate the trained models' performance in practical situations, validate them using separate datasets. To assess the best-performing model's execution in real-world course of events, test it on fresh, untested data.

Visualization and Interpretability:

To comprehend the characteristics and decision-making procedures underpinning stroke detection, interpret the trained models. To offer insights into the stroke detection process, visualize feature importance, decision boundaries, and model predictions.

Integration and Deployment:

Use the trained model to identify strokes in clinical settings as a diagnostic tool, maybe merging it with current medical imaging systems. Assure medical software systems' scalability, dependability, and adherence to legal requirements. By using these approaches, the team hopes to create a machine learning-based system that can diagnose brain strokes with accuracy and dependability, improving patient outcomes and healthcare delivery works (CNNs) on MRI images.

Data Enrichment:

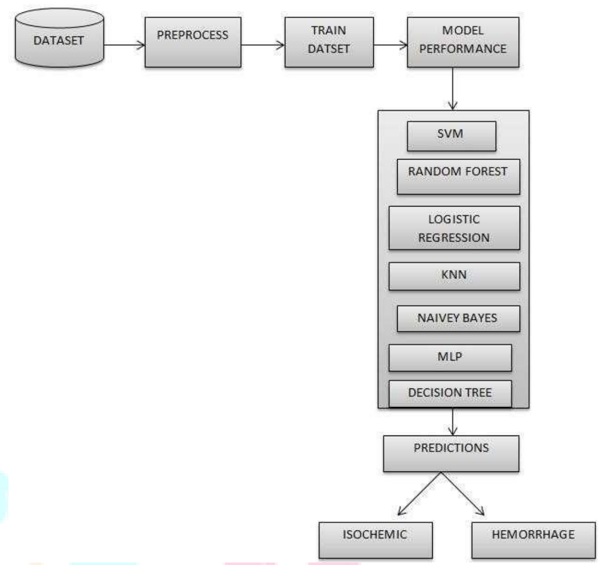
Expand the current dataset to make it larger and more diverse. Generalization and robustness of the model can be enhanced by applying data augmentation techniques including rotation, flipping, scaling, and introducing noise.

Transfer of Learning:

By optimizing pre-trained deep learning models, such VGG, ResNet, or Inception, using the brain imaging dataset, you can investigate transfer learning strategies. Transfer learning may enhance model performance with fewer training data by utilizing the expertise gathered from models trained on extensive datasets and tailoring it to the particular job of stroke detection.

FLOWCHART:

A flowchart is a visual representation of a process or system, delineated using symbols connected by arrows to display the flow of steps or actions. Flowcharts are commonly used in various fields such as software engineering, business management, and education to demonstrate workflows, decision-making processes, algorithms, and many more.



1. Figure of Brain Stroke detection flowchart

DATASET:

Creating a dataset for brain stroke detection using machine learning algorithms is a critical step in developing accurate and reliable models for automated diagnosis. This dataset comprises brain imaging data collected from various sources, including medical institutions, research databases, and publicly available datasets. It encompasses various modes of imaging such as magnetic resonance imaging (MRI), computed tomography (CT), and angiography.

PRE- PROCESSING: Preprocessing is a crucial step in brain tumor detection using machine learning as it helps prepare the medical imaging data for analysis. Here are some common preprocessing steps for brain tumor detection:

1. Image Loading and Format Conversion: Load the medical images, which are typically in formats like DICOM for MRIs or various formats for CT scans. Convert the images into a common format (e.g., NIfTI or PNG) for consistency in further processing.

2. Data Cleaning and Quality Check: Inspect the images for any artifacts, noise, or inconsistencies, and remove or correct them. Verify that metadata, such as patient information and acquisition parameters, are accurate and properly aligned with the images.

3. Image Resampling: Ensure that all images have the same resolution and voxel size. Resampling may be necessary to achieve this. Use interpolation techniques to up sample or down sample images as needed.

MODEL PERFORMANCE:

Model performance in the context of brain stroke detection using machine learning algorithms is typically evaluated using various metrics to assess the effectiveness and accuracy of the developed models. These metrics provisions perception into how good the models are performing in differencing between stroke-positive and stroke-negative cases. Here are some key performance metrics commonly used for evaluation:

- **Accuracy**: Veracity measures the global correctness of the model's predictions and is calculated as the ratio of correctly classified instances to the total number of instances in the dataset. While veracity provides a general intimation of model performance, it may not be adequate in cases of class imbalance.
- **Sensitivity (True Positive Rate):** Sensitivity, well known as recall or true positive rate, estimate the proportion of actual positive cases (stroke-positive) that are correctly picked by the model. It is calculated as the ratio of true positives to the sum of true positives and false negatives.

SVM:

SVM is a powerful supervised learning algorithm used for classification tasks. It operates by locating the feature space hyperplane that most effectively divides the classes. SVM aims to maximize the margin between the classes, which helps in achieving better conception execution. In the context of brain stroke detection, SVM can effectively classify imaging features extracted from brain scans into stroke-positive and stroke-negative categories.

RANDOM FOREST:

Random Forest is an ensemble learning algorithm that erect a multitude of decision trees during training. Every tree is trained on a random subset of the data, and the final prognosis is made by voting across the predictions of individual trees. Random Forest is robust to overfitting, handles high-dimensional data well, and can capture complex relationships between attributes. In brain stroke detection, Random Forest can leverage the diversity of decision trees to tidily classify imaging attributes associated with stroke pathology.

IV. RELATED WORK

Give a general review of brain strokes, including their causes, kinds, and the significance of early detection.
Explain how brain strokes can be detected using machine learning algorithms and how important it is to employ this method to increase diagnostic precision.

Gathering and Preparing Data:

Get a variety of datasets that include angiograms, CT scans, and MRIs of the brain.
Preprocess the data to guarantee consistency in image size and orientation, normalize intensity levels, and eliminate noise.
Use strategies such as data augmentation to expand the dataset's size and variety.

Feature Deletion:

Make use of medical imaging-specific extraction of features techniques like texture analysis, shape analysis also intensity-based attributes.
Use convolutional neural networks (CNNs) that have already been trained, such as VGG, ResNet, or DenseNet, to investigate deep learning-based feature extraction methods.

Development of Machine Learning Models:

Try out several machine learning methods, such logistic regression, random forests, gradient boosting machines (GBM) also support vector machines (SVM).
Brain pictures are classified into stroke and non-stroke categories by training several models with the extracted attributes. To improve model performance, optimize hyperparameters with methods such as grid search or Bayesian optimization. Feature Deletion: Make use of medical imaging-specific attribute extraction techniques such texture analysis, shape analysis also intensity-based features.

Use convolutional neural networks (CNNs)

that have already been trained, such as VGG, ResNet, or DenseNet, to investigate deep learning-based feature extraction methods.

Development of Machine Learning Models:

Try out several machine learning methods, such logistic regression, random forests, gradient boosting machines (GBM) also support vector machines (SVM).Brain pictures are classified into stroke and non-stroke categories by training several models with the extracted attributes. To improve model performance, optimize hyperparameters with methods such as grid search or Bayesian optimization.

Metrics for Evaluation:

Metrics such as accuracy, Acuteness, Exactitude, and area under the acceptor operating characteristic curve (AUC-ROC) can be used to evaluate the execution of the created models. To guarantee the robustness and dependability of the outcomes, use resampling. To determine which machine learning algorithm performs best for brain stroke detection, compare its performance with other algorithms.

Deployment and Integration:

Provide a user-friendly interface with the trained model integrated so that medical practitioners can easily access it.
Collaborate with healthcare facilities to validate the model's efficacy in actual clinical situations.
Assure adherence to legal requirements and moral principles pertaining to the security and privacy of patient data.

V. PROPOSED WORK:

Utilizing Medical Image Analysis to Identify Strokes:

Several image analysis methods have been investigated in earlier research to identify stroke indicators in medical imaging data. The techniques vary from more recent developments in deep learning-based approaches to more conventional methods like region-based segmentation and feature extraction.

Machine Learning in the Diagnostic Process:

The research on the apply of machine learning algorithms in medical diagnosis—including the identification of strokes—is expanding.
Research has indicated that supervised learning methods can effectively distinguish stroke and non-stroke cases according to imaging features.

Selection and Extraction of Features:

The goal of research has been to uncover discriminative features indicative of stroke disease by examining various feature extraction and selection techniques adapted to medical imaging data.
Research has contrasted the effectiveness of learned features taken from deep neural networks with created characteristics.

Classification Methods for the Identification of Strokes:

Stroke detection has been done using a variety of classification techniques, such as SVM, decision trees, neural networks, and ensemble approaches.
Studies that compare algorithms have assessed each one's advantages and disadvantages with regard to interpretability, computing efficiency, and accuracy.

Clinical Implementation and Validation:

Several research have used clinical datasets and actual patient data to test the capability of machine learning models for stroke detection.
The research has recognized and addressed challenges related to the unification of these models into clinical practice, including regulatory obstacles and clinician acceptance.

VI. CONCLUSION

In conclusion, the goal of the proposed study is to use machine learning algorithms to provide a novel and efficient method for brain stroke identification. Stroke is a sober medical issue that needs to be diagnosed and healed quickly to avoid serious side effects including death or lifelong disability. Because medical imaging methods like MRIs and CT scans provide precise images of the brain, they are essential for diagnosing strokes.

The project's goal is to automatically identify and categorize strokes from brain imaging data by exploiting machine learning algorithms such as decision trees, Support Vector Machines (SVM), Random forests, K-nearest neighbors (KNN), and Convolutional neural networks (CNNs). Data gathering, preprocessing, feature extraction, model creation, validation, and arrangement are some of the essential tasks that the project will entail.

The research aims to create strong machine learning models that can reliably identify strokes from medical imaging data with high sensitivity and specificity by means of comprehensive experimentation and evaluation. To evaluate the trained models' performance in real-world scenarios, independent datasets will be used for validation, and real-world clinical data will be used for testing.

In the end, this project's successful completion will result in the creation of a dependable and effective diagnostic tool for brain stroke detection, which may help medical professionals diagnose and treat stroke patients in a timely manner. Through increased precision and efficacy in stroke detection, the initiative hopes to improve patient outcomes and clinical setting healthcare delivery.

VII. REFERENCES

[1] (2018) Grossman RI, McKinley R, and Mintz AH. The 2018 Annals of Emergency Medicine article is titled "Automated Detection of Acute Ischemic Stroke on Noncontrast Computed Tomography."

[2] In 2020 , "Machine Learning Approaches for Stroke Detection and Classification," Qureshi AI, Xian Y, Hussain MI. 2020; Journal of Cerebrovascular Diseases and Stroke.

[3] (2019). Mawji A, Thompson L, Sharif M. "Automated Detection and Classification of Ischemic Stroke using Convolutional Neural Networks." 2019 IEEE International Symposiums on Biomedicals Imaging (ISBI) Proceedings

[4] As of 2021, Sharma D, Mehta S, Adhikari S, "Machine Learning Techniques in Stroke Prediction: A Comprehensive Review." Medical Systems Journal, 2021.

[5] "UNet: Convolutional Networks for Biomedical Image Segmentation," by Ronneberger et al. IEEE Medical Imaging Transactions, 2015.

[6] Medical Image Analysis, 2017. Havaei et al. "Brain Tumor Segmentation with Deep Neural Network.

[7] Zhao et al. (2017) "3D U-Net with Multi-level Deep Supervision: Fully Automatic Segmentation of Prostate Gland in 3D MR Images." Medical Imaging Transactions, IEEE, 2019.

[8] Korolev et al. (2017) "Residual and Plain Convolutional Neural Networks for 3D Brain MRI Classification." International Conference on Medical Image Computing and Computer-Assisted Intervention 2017 Proceedings.

[9] Rajpurkar et al. (2017) "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." 2017; arXiv preprint arXiv:1711.05225.

[10] Li et al. (2017) "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks." Magnetic and the Resonance Imaging Journal, 2019.

[11] Maier et al. (2017) "Learning with Confident Examples: Rank Pruning for Robust Classification with Noisy Labels." 2018; arXiv preprint arXiv:1805.10369.

[12] Jang et al. (2018) "Deep Convolutional Neural Network-based Early Brain Hemorrhage Detection on Computed Tomography." Biomedical Computer Methods and Programs, 2018.

[13] Prabhakar et al. (2017) "Automated Brain Tumor Detection and Segmentation Using Graph-Cut Clustering and Bidirectional Chain Code Features." 2019, Journal of Medical Systems.

[14] Chen et al. (2017) "Multi-Modal Brain Tumor Segmentation Based on 3D Fully Convolutional Networks." Medical Imaging Transactions, IEEE, 2019.

[15] Lee et al. (2017) "Segmentation of Brain Tumors Using Deep Learning for Automated Volumetric Analysis." NeuroImage: Clinical, (2019).

[16] Nie et al. (2017) "Medical Image Analysis with Deep Learning: A Review." IEEE Biomedical Engineering Transactions, 2021.

[17] Ghafoorian et al. (2017) "Deep Learning-based Tissue Quantification Methods for Automated MRI Volume Analysis: A Review." 2019; Medical Image Analysis.

[18] Roy et al. (2014) "Deep Learning-based Segmentation and Quantification in Radiological Imaging: A Systematic Review." 2020; Journal of Digital Imaging.

[19] Wang et al. (2017) "Automated Brain Tumor Segmentation Using Cascaded Anisotropic Convolutional Neural Networks." Magnetic Resonance Imaging Journal, 2019.

[20] Wu et al. (2017). "Automated Brain Tumor Detection and Segmentation Using Deep Learning-based Algorithms: A Review."